RAW SEQUENCE LISTING ERROR REPORT
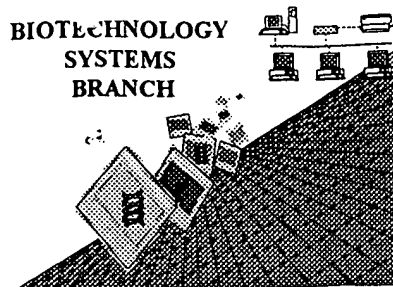
The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: _09/836,911_

Source: _OIPE_

Date Processed by STIC: _5/3/2001_

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov **or phone 703-306-4119 (R. Wax)**
PATENTIN 3.0 e-mail help: patin3help@uspto.gov **or phone 703-306-4119 (R. Wax)**

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
· **http://www.uspto.gov/web/offices/pac/checker**

```
        RAW SEQUENCE LISTING              DATE: 05/03/2001
        PATENT APPLICATION: US/09/836,911  TIME: 15:56:14

        Input Set : A:\402iseq.001
        Output Set: N:\CRF3\05032001\I836911.raw
```

Does Not Comply
Corrected Diskette Needed

P S

```
        SEQUENCE LISTING
C--> 4 (1) GENERAL INFORMATION:
     6       (i) APPLICANT: Hadlaczky, Gyula
     7                      Szalay, Aladar
C--> 9      (ii) TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
     10                        AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
     12     (iii) NUMBER OF SEQUENCES: 34
     14      (iv) CORRESPONDENCE ADDRESS:
     15           (A) ADDRESSEE: Heller Ehrman White & McAuliffe
     16           (B) STREET: 4350 La Jolla Village Drive, 6th Floor
     17           (C) CITY: San Diego
     18           (D) STATE: CA
     19           (E) COUNTRY: USA
     20           (F) ZIP: 92122
     22       (v) COMPUTER READABLE FORM:
     23           (A) MEDIUM TYPE: Diskette
     24           (B) COMPUTER: IBM Compatible
     25           (C) OPERATING SYSTEM: DOS
     26           (D) SOFTWARE: FastSEQ Version 1.5
     28      (vi) CURRENT APPLICATION DATA:
C--> 29           (A) APPLICATION NUMBER: US/09/836,911
C--> 30           (B) FILING DATE: 17-Apr-2001
     50           (C) CLASSIFICATION:
C--> 47     (vii) PRIOR APPLICATION DATA:
     33           (A) APPLICATION NUMBER: 08/835,682
     34           (B) FILING DATE: 10-APR-1997
     38           (A) APPLICATION NUMBER: 08/695,191
     39           (B) FILING DATE: 07-AUG-1996
     43           (A) APPLICATION NUMBER: 08/682,080
     44           (B) FILING DATE: 15-JUL-1996
     48           (A) APPLICATION NUMBER: 08/629,822
     49           (B) FILING DATE: 10-APR-1996
     52    (viii) ATTORNEY/AGENT INFORMATION:
     53           (A) NAME: Seidman, Stephanie L
     54           (B) REGISTRATION NUMBER: 33,779
     55           (C) REFERENCE/DOCKET NUMBER: 24601-402I
     58      (ix) TELECOMMUNICATION INFORMATION:
     59           (A) TELEPHONE: 858-450-8403
     60           (B) TELEFAX: 858-587-5360
     61           (C) TELEX:
     63 (2) INFORMATION FOR SEQ ID NO: 1:
     65       (i) SEQUENCE CHARACTERISTICS:
     66           (A) LENGTH: 1293 base pairs
     67           (B) TYPE: nucleic acid
     68           (C) STRANDEDNESS: single
     69           (D) TOPOLOGY: linear
     71      (ii) MOLECULE TYPE: Genomic DNA
```

RAW SEQUENCE LISTING   DATE: 05/03/2001
PATENT APPLICATION: US/09/836,911   TIME: 15:56:14

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
        72    (iii) HYPOTHETICAL: NO
C--> 73    (iv) ANTI-SENSE: NO
W--> 74    (v) FRAGMENT TYPE:
        75    (vi) ORIGINAL SOURCE:
        76    (ix) FEATURE:
        78    (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
        80    GAATTCATCA TTTTTCANGT CCTCAAGTGG ATGTTTCTCA TTTNCCATGA TTTTAAGTTT      60
        81    TCTCGCCATA TTCCTGGTCC TACAGTGTGC ATTTCTCCAT TTTNCACGTT TTNCAGTGAT     120
        82    TTCGTCATTT TCAAGTCCTC AAGTGGATGT TTCTCATTTN CCATGAATTT CAGTTTTCTN     180
        83    GCCATATTCC ACGTCCTACA GNGGACATTT CTAAATTTNC CACCTTTTTC AGTTTTCCTC     240
        84    GCCATATTTC ACGTCCTAAA ATGTGTATTT CTCGTTTNCC GTGATTTTCA GTTTTCTCGC     300
        85    CAGATTCCAG GTCCTATAAT GTGCATTTCT CATTTNNCAC GTTTTTCAGT GATTTCGTCA     360
        86    TTTTTTCAAG TCGGCAAGTG GATGTTTCTC ATTTNCCATG ATTTNCAGTT TTCTTGNAAT     420
        87    ATTCCATGTC CTACAATGAT CATTTTTAAT TTTCCACCTT TTCATTTTTC CACGCCATAT     480
        88    TTCATGTCCT AAAGTGTATA TTTCTCCTTT TCCGCGATTT TCAGTTTTCT CGCCATATTC     540
        89    CAGGTCCTAC AGTGTGCATT CCTCATTTTT CACCTTTTTC ACTGATTTCG TCATTTTTCA     600
        90    AGTCGTCAAC TGGATCTTTC TAATTTTCCA TGATTTTCAG TTATCTTGTC ATATTCCATG     660
        91    TCCTACAGTG GACATTTCTA AATTTTCCAA CTTTTTCAAT TTTTCTCGAC ATATTTGACG     720
        92    TGCTAAAGTG TGTATTTCTT ATTTTCCGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGTC     780
        93    CTAATAGTGT GCATTTCTCA TTTTTCACGT TTTTCAGTGA TTTCGTCATT TTTTCCAGTT     840
        94    GTCAAGGGGA TGTTTCTCAT TTTCCATGAG TGTCAGTTTT CTTGCTATAT TCCATGTCCT     900
        95    ACAGTGACAT TTCTAAATAT TATACCTTTT TCAGTTTTTC TCACCATATT TCACGTCCTA     960
        96    AAGTATATAT TTCTCATTTT CCCTGATTTT CAGTTCCTT GCCATATTCC AGGTCCTACA    1020
        97    GTGTGCATTT CTCATTTTTC ACGTTTTTCA GTAATTTCTT CATTTTTTAA GCCCTCAAAT    1080
        98    GGATGTTTCT CATTTTCCAT GATTTTCAGT TTTCTTGCCA TATACCATGT CCTACAGTGG    1140
        99    ACATTTCTAA ATTATCCACC TTTTTCAGTT TTTCATCGGC ACATTTCACG TCCTAAAGTG    1200
       100    TGTATTTCTA ATTTTCAGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGAC CTACAGTGTG    1260
       101    CATTTCTCAT TTTTCACGTT TTTCAGTGAA TTC                                1293
       103    (2) INFORMATION FOR SEQ ID NO: 2:
       105        (i) SEQUENCE CHARACTERISTICS:
       106            (A) LENGTH: 1044 base pairs
       107            (B) TYPE: nucleic acid
       108            (C) STRANDEDNESS: single
       109            (D) TOPOLOGY: linear
       111        (ii) MOLECULE TYPE: Genomic DNA
       112        (iii) HYPOTHETICAL: NO
C--> 113    (iv) ANTI-SENSE: NO
W--> 114    (v) FRAGMENT TYPE:
       115        (vi) ORIGINAL SOURCE:
       116        (ix) FEATURE:
       118        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
       120    AGGCCTATGG TGAAAAAGGA AATATCTTCC CCTGAAAACT AGACAGAAGG ATTCTCAGAA      60
       121    TCTTATTTGT GATGTGCGCC CCTCAACTAA CAGTGTTGAA GCTTTCTTTT GATAGAGCAG     120
       122    TTTTGAAACA CTCTTTTTGT AAAATCTGCA AGAGGATATT TGGATAGCTT TGAGGATTTC     180
       123    CGTTGGAAAC GGGATTGTCT TCATATAAAC CCTAGACAGA AGCATTCTCA GAAGCTTCAT     240
       124    TGGGATGTTT CAGTTGAAGT CACAGTGTTG AACAGTCCCC TTTCATAGAG CAGGTTTGAA     300
       125    ACACTCTTTT TTGTAGTATC TGGAAGTGGA CATTTGGAGC GATCTCAGGA CTGCGGTGAA     360
       126    AAAGGAAATA TCTTCCAATA AAAGCTAGAT AGAGGCAATG TCAGAAACCT TTTTCATGAT     420
       127    GTATCTACTC AGCTAACAGA GTTGAACCTT CCTTTGAGAG AGCAGTTTTG AAACACTCTT     480
```

RAW SEQUENCE LISTING   DATE: 05/03/2001
PATENT APPLICATION: US/09/836,911   TIME: 15:56:14

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
        128  TTTGTGGAAT CTGCAAGTGG ATATTTGTCT AGCTTTGAGG ATTTCGTTGG GAAACGGGAT    540
        129  TACATATAAA AAGCAGACAG CAGCATTCCC AGAAACTTCT TTGTGATGTT TGCATTCAAG    600
        130  TCACAGAGTT GAACATTCCC TTTCATAGAG CAGGTTTGAA ACACACTTTT TGATGTATCT    660
        131  GGATGTGGAC ATTTGCAGCG CTTTCAGGCC TAAGGTGAAA AGGAAATATC TTCCCCTGAA    720
        132  AACTAGACAG AAGCATTCTC AGAAACTTAT TTGTGATGTG CGCCCTCAAC TAACAGTGTT    780
        133  GAAGCTTTCT TTTGATAGAG GCAGTTTTGA AACACTCTTT TGTGGAATCT GCAAGTGGAT    840
        134  ATTTGTCTAG CTTTGAGGAT TTCTTTGGAA ACGGGATTAC ATATAAAAAG CAGACAGCAG    900
        135  CATTCCCAGA ATCTTGTTTG TGATGTTTGC ATTCAAGTCA CAGAGTTGAA CATTCCCTTT    960
        136  CAGAGAGCAG GTTTGAACAC TCTTTTTATA GTATCTGGAT GTGGACATTT GGAGCGCTTT   1020
        137  CAGGGGGGAT CCTCTAGAAT TCCT                                          1044
        141  (2) INFORMATION FOR SEQ ID NO: 3:
        143       (i) SEQUENCE CHARACTERISTICS:
        144           (A) LENGTH: 2492 base pairs
        145           (B) TYPE: nucleic acid
        146           (C) STRANDEDNESS: single
        147           (D) TOPOLOGY: linear
        149       (ii) MOLECULE TYPE: Genomic DNA
        150      (iii) HYPOTHETICAL: NO
C--> 151       (iv) ANTI-SENSE: NO
W--> 152        (v) FRAGMENT TYPE:
        153       (vi) ORIGINAL SOURCE:
        154       (ix) FEATURE:
        156       (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
        158  CTGCAGCTGG GGGTCTCCAA TCAGGCAGGG GCCCCTTACT ACTCAGATGG GGTGGCCGAG     60
        159  TAGGGGAAGG GGGTGCAGGC TGCATGAGTG GACACAGCTG TAGGACTACC TGGGGGCTGT    120
        160  GGATCTATGG GGGTGGGGAG AAGCCCAGTG ACAGTGCCTA GAAGAGACAA GGTGGCCTGA    180
        161  GAGGGTCTGA GGAACATAGA GCTGGCCATG TTGGGGCCAG GTCTCAAGCA GGAAGTGAGG    240
        162  AATGGGACAG GCTTGAGGAT ACTCTACTCA GTAGCCAGGA TAGCAAGGAG GGCTTGGGGT    300
        163  TGCTATCCTG GGGTTCAACC CCCCAGGTTG AAGGCCCTGG GGGAGATGGT CCCAGGACAT    360
        164  ATTACAATGG ACACAGGAGG TTGGGACACC TGGAGTCACC AAACAAAACC ATGCCAAGAG    420
        165  AGACCATGAG TAGGGGTGTC CAGTCCAGCC CTCTGACTGA GCTGCATTGT TCAAATCCAA    480
        166  AGGGCCCCTG CTGCCACCTA GTGGCTGATG GCATCCACAT GACCCTGGGC CACACGCGTT    540
        167  TAGGGTCTCT GTGAAGACCA AGATCCTTGT TACATTGAAC GACTCCTAAA TGAGCAGAGA    600
        168  TTTCCACCTA TTCGAAACAA TCACATAAAA TCCATCCTGG AAAAAGCCTG GGGGATGGCA    660
        169  CTAAGGCTAG GGATAGGGTG GGATGAAGAT TATAGTTACA GTAAGGGGTT TAGGGTTAGG    720
        170  GATCAACGTT GGTTAGGAGT TAGGGATACA GTAGGGTACC GGTAGGGTTA GGGGTTAGGG    780
        171  TTAGGGGTTA GGGTTAGGGT TAGGGTTAGG GTTAGGGTTA GGGGTTAGGG GTTAGGGTTA    840
        172  GGGTTAGGTT TTGGGGTGGC GTATTTTGGT CTTATACGCT GTGTTCCACT GGCAATGAAA    900
        173  AGAGTTCTTG TTTTTCCTTC AGCAATTTGT CATTTTTAAA AGAGTTTAGC AATTCTAACA    960
        174  GATATAGACC AGCTGTGCTA TCTCATTGTG GTTTTCAATT GTAACCACAT TGTGGTTTCA   1020
        175  ATGTGTTTAC TTGCCATCTG TAGATCTTCT TTGCGTGAGG TGTCTGTTCA GATGTGTGTG   1080
        176  CATTTCTTGN NTTTNGGCTG TTTAACTTAT TGTTTAGTTT TAATAATTTT TTATATATTT   1140
        177  GAAGACAAAT CTTTCTCAGA TGTGTATTTG CAAATATTTC TTCAATATGA GGCTTGCTTT   1200
        178  TGTCTCTAAC AAGGTCTCTT CAGAGATAAC TTAAATATAA GAAATCCACA CTGTCACTTC   1260
        179  TTTTGTGTAT ATCTACCTTT TGTGTCATTT GTTAAAATTC ATTACCAAAC CCAAAGGCAG   1320
        180  ATAGCTTTTC TTCTATTGTT TCTTCTAGAA ATTTGTATAG TTTTGCATTT TTAGTGTAAG   1380
        181  GATGATTTTG AGTGATTATT TGTGTAAGTT GTAAAGTTTT CGTCTATATC CATATCATTT   1440
        182  CTTATGGTTT CCAATTAATC GTTCCCTCAC TATTTTTGGG AAAGACACAG GATAGTGGGC   1500
        183  TTTGTTAGAG TAGATAGGTA GCTAGACATG AACAGGAGGG GGCCTCCTGG AAAAGGGAAA   1560
```

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
184  GTCTGGGAAG GCTCACCTGG AGGACCACCA AAAATTCACA TATTAGTAGC ATCTCTAGTG   1620
185  CTGGAGTGGA TGGGCACTTG TCAATTGTGG GTAGGAGGGA AAAGAGGTCC TATGCAGAAA   1680
186  GAAACTCCCT AGAACTCCTC TGAAGATGCC CCAATCATTC ACTCTGCAAT AAAAATGTCA   1740
187  GAATATTGCT AGCTACATGC TGATAAGGNN AAAGGGGACA TTCTTAAGTG AAACCTGGCA   1800
188  CCATAAGTAC AGATTAGGGC AGAGAAGGAC ATTCAAAAGA GGCAGGCGCA GTAGGTACAA   1860
189  ACGTGATCGC TGTCAGTGTG CCTGGGATGG CGGGAAGGAG GCTGGTGCCA GAGTGGATTC   1920
190  GTATTGATCA CCACACATAT ACCTCAACCA ACAGTGAGGA GGTCCCACAA GCCTAAGTGG   1980
191  GGCAAGTTGG GGAGCTAAGG CAGTAGCAGG AAAACCAGAC AAAGAAAACA GGTGGAGACT   2040
192  TGAGACAGAG GCAGGAATGT GAAGAAATCC AAAATAAAAT TCCCTGCACA GGACTCTTAG   2100
193  GCTGTTTAAT GCATCGCTCA GTCCCACTCC TCCCTATTTT TCTACAATAA ACTCTTTACA   2160
194  CTGTGTTTCT TTTCAATGAA GTTATCTGCC ATCTTTGTAT TGCCTCTTGG TGAAAATGTT   2220
195  TCTTCCAAGT TAAACAAGAA CTGGGACATC AGCTCTCCCC AGTAATAGCT CCGTTTCAGT   2280
196  TTGAATTTAC AGAACTGATG GGCTTAATAA CTGGCGCTCT GACTTTAGTG GTGCAGGAGG   2340
197  CCGTCACACC GGGACCAAGA GTGCCCTGCC TAGTCCCCAT CTGCCCGCAG GTGGCGGCTG   2400
198  CCTCGACACT GACAGCAATA GGGTCCGGCA GTGTCCCCAG CTGCCAGCAG GGGGCGTACG   2460
199  ACGACTACAC TGTGAGCAAG AGGGCCCTGC AG                                 2492
201  (2) INFORMATION FOR SEQ ID NO: 4:
203       (i) SEQUENCE CHARACTERISTICS:
204            (A) LENGTH: 28 base pairs
205            (B) TYPE: nucleic acid
206            (C) STRANDEDNESS: single
207            (D) TOPOLOGY: linear
209      (ii) MOLECULE TYPE: Genomic DNA
210     (iii) HYPOTHETICAL: NO
```
C--> 211   **(iv) ANTI-SENSE: NO**
W--> 212   **(v) FRAGMENT TYPE:**
```
213      (vi) ORIGINAL SOURCE:
214      (ix) FEATURE:
216      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
218  GGGGAATTCA TTGGGATGTT TCAGTTGA                                      28
220  (2) INFORMATION FOR SEQ ID NO: 5:
222       (i) SEQUENCE CHARACTERISTICS:
223            (A) LENGTH: 29 base pairs
224            (B) TYPE: nucleic acid
225            (C) STRANDEDNESS: single
226            (D) TOPOLOGY: linear
228      (ii) MOLECULE TYPE: Genomic DNA
229     (iii) HYPOTHETICAL: NO
```
C--> 230   **(iv) ANTI-SENSE: NO**
W--> 231   **(v) FRAGMENT TYPE:**
```
232      (vi) ORIGINAL SOURCE:
233      (ix) FEATURE:
235      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
237  CGAAAGTCCC CCCTAGGAGA TCTTAAGGA                                     29
239  (2) INFORMATION FOR SEQ ID NO: 6:
241       (i) SEQUENCE CHARACTERISTICS:
242            (A) LENGTH: 47 base pairs
243            (B) TYPE: nucleic acid
244            (C) STRANDEDNESS: single
```

RAW SEQUENCE LISTING   DATE: 05/03/2001
PATENT APPLICATION: US/09/836,911   TIME: 15:56:14

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
        245            (D) TOPOLOGY: linear
W--> 247       (ii) MOLECULE TYPE: DNA
        248       (iii) HYPOTHETICAL: NO
C--> 249        (iv) ANTI-SENSE: NO
W--> 250         (v) FRAGMENT TYPE:
        251        (vi) ORIGINAL SOURCE:
        252        (ix) FEATURE:
        254        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
        256 CCGCTTAATA CTCTGATGAG TCCGTGAGGA CGAAACGCTC TCGCACC          47
        260 (2) INFORMATION FOR SEQ ID NO: 7:
        262        (i) SEQUENCE CHARACTERISTICS:
        263            (A) LENGTH: 25 base pairs
        264            (B) TYPE: nucleic acid
        265            (C) STRANDEDNESS: single
        266            (D) TOPOLOGY: linear
        268       (ii) MOLECULE TYPE: Genomic DNA
        269       (iii) HYPOTHETICAL: NO
C--> 270        (iv) ANTI-SENSE: NO
W--> 271         (v) FRAGMENT TYPE:
        272        (vi) ORIGINAL SOURCE:
        273        (ix) FEATURE:
        275        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
        277 CGATTTAAAT TAATTAAGCC CGGGC                                  25
        280 (2) INFORMATION FOR SEQ ID NO: 8:
        282        (i) SEQUENCE CHARACTERISTICS:
        283            (A) LENGTH: 27 base pairs
        284            (B) TYPE: nucleic acid
        285            (C) STRANDEDNESS: single
        286            (D) TOPOLOGY: linear
        288       (ii) MOLECULE TYPE: Genomic DNA
        289       (iii) HYPOTHETICAL: NO
C--> 290        (iv) ANTI-SENSE: NO
W--> 291         (v) FRAGMENT TYPE:
        292        (vi) ORIGINAL SOURCE:
        293        (ix) FEATURE:
        295        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
        297 TAAATTTAAT TAATTCGGGC CCGTCGA                                27
        299 (2) INFORMATION FOR SEQ ID NO: 9:
        301        (i) SEQUENCE CHARACTERISTICS:
        302            (A) LENGTH: 69 base pairs
        303            (B) TYPE: nucleic acid
        304            (C) STRANDEDNESS: single
        305            (D) TOPOLOGY: linear
        307       (ii) MOLECULE TYPE: Genomic DNA
        310        (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
        312 ATG TAC AGG ATG CAA CTC CTG TCT TGC ATT GCA CTA AGT CTT GCA CTT    48
        313 Met Tyr Arg Met Gln Leu Leu Ser Cys Ile Ala Leu Ser Leu Ala Leu
W--> 315 GTC ACA AAC AGT GCA CCT ACT                                   69
        316 Val Thr Asn Ser Ala Pro Thr
```

(Per Sequence Rules)

) Number the amino acids under every 5 amino acids — do not use TAB codes between numbers 5/3/01

Please edit all subsequent sequences containing these errors

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
L:4   M:220 C: Keyword misspelled or invalid format, [(1) GENERAL INFORMATION:]
L:9   M:220 C: Keyword misspelled or invalid format, [(ii) TITLE OF INVENTION:]
L:29  M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30  M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:37  M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:42  M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:47  M:220 C: Keyword misspelled or invalid format, [(vii) PRIOR APPLICATION DATA:]
L:73  M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:78  M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=1
L:74  M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=1
L:113 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:118 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=2
L:114 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=2
L:151 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:156 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=3
L:152 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=3
L:211 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:216 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=4
L:212 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=4
L:230 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:235 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=5
L:231 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=5
L:249 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:254 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=6
L:247 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:250 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=6
L:270 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:275 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=7
L:271 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=7
L:290 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:295 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=8
L:291 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=8
L:308 M:220 C: Keyword misspelled or invalid format, [(D) OTHER INFORMATION:]
L:308 M:220 C: Keyword misspelled or invalid format, Poss data loss, Seq 9, (D) OTHER INFORMATION:
L:315 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:9
L:339 M:220 C: Keyword misspelled or invalid format, [(H) DOCUMENT NUMBER:]
L:433 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:438 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=11
L:434 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=11
L:452 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:457 M:246 W: Invalid value of Alpha Sequence Header Field, [FEATURE:], SeqNo=12
L:453 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=12
L:471 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:472 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=13
L:512 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:513 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=14
L:553 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:554 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=15
```

Input Set : A:\402iseq.001
Output Set: N:\CRF3\05032001\I836911.raw

```
L:593  M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:594  M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=16
L:979  M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:980  M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=17
L:1713 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1714 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=18
L:1734 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1735 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=19
L:1764 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1765 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=20
L:1789 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1790 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=21
L:1813 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1814 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=22
L:1837 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1838 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=23
L:1866 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1867 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=24
L:1896 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1897 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=25
L:1914 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1915 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=26
L:1932 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1933 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=27
L:1950 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1951 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=28
L:1968 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1969 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=29
L:1987 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:1988 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=30
L:2005 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:2006 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=31
L:2023 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:2024 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=32
L:2041 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:2042 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=33
L:2059 M:220 C: Keyword misspelled or invalid format, [(iv) ANTI-SENSE:]
L:2060 M:246 W: Invalid value of Alpha Sequence Header Field, [FRAGMENT TYPE:], SeqNo=34
```